# STRESS-STRENGTH MODELLING: RANKED SET AND SIMPLE RANDOM SAMPLING IN GENERALIZED INVERSE WEIBULL ANALYSIS

Surinder Kumar, Bhupendra Meena, Rahul Shukla, Shivendra Pratap Singh
•

Department of Statistics,
Babasaheb Bhimrao Ambedkar University, 226025, U.P, India
surinderntls@gmail.com
Corresponding Author: bhupendrakm57@gmail.com
rahul.shukla.stats@gmail.com
shivendra15.07@gmail.com

**Abstract**

*This study explores the stress-strength reliability model (P) for Generalized Inverse Weibull (GIW) distribution through transformation techniques. We compare two sampling methods: ranked set sampling (RSS) and simple random sampling (SRS), where stress and strength are two independent random variables from the GIW distribution respectively. RSS, is used for estimating stress-strength model, as this technique of sampling is more efficient alternative of SRS for obtaining the more informative sample. In this article, the maximum likelihood estimator (MLE) for stress-strength model is obtained through transforming technique. MLE estimates of stress-strength obtained through Ranked set sampling (RSS) methods are evaluated against corresponding estimates derived from simple random sampling (SRS) to understand their relative effectiveness and accuracy. The statistical estimators derived from Ranked Set Sampling (RSS) methodology exhibit superior efficiency relative to their Simple Random Sampling (SRS) counterparts. The empirical utility of RSS-based estimation procedures is subsequently validated through application to real datasets.*

**Keywords**: Stress-strength reliability, simple random sampling, ranked set sampling, generalized inverse Weibull distribution, maximum likelihood estimation.

## 1. Introduction

The stress-strength model is a fundamental concept in reliability engineering and statistics. It is used to assess the probability of failure or success in a system subject to the random variations in stress and strength. This model is employed by many researchers in various fields, including engineering, materials science, quality control, and finance etc. The probability that a system's random stress Y is less than its random strength X is represented by $= P_r(Y < X)$ in the context of stress-strength model. In other words, it calculates the probability of failure in the stress-strength model. The system failure occurs when the stress exceeds the strength. Recently, the problem of stress-strength model is evaluated by an alternative approach of sampling proposed by Mclntyre [1] The pioneering investigations of Birnbaum [2] and Birnbaum and McCarty [3] represent the initial academic exploration of this fundamental problem. Church and Harris [4] were the first to use the phrase "stress-strength". Since then, a sizable amount of work has been completed from both a parametric and non-parametric perspective. For earlier bibliography one may refers to, Chaturvedi and Kumar [5], Kotz et al. [6], Kundu and Gupta [7][8], Raqab and Kundu [9], Kundu and Raqab [10], Krishnamoorthy et al. [11], Hassan [12], Wang et al. [13], Kayal et al. [14], Kumar and Chaturvedi [15]. In the above referred studies the estimation for the considered model is based on SRS.

The circumstances in which it is challenging to take the actual measurement for sample units (costly, destructive, time consuming), the RSS strategy can be used in under these circumstances which maintains the accuracy of our statistical judgements and reduces the sample size. Akgul and Senoglu [16] obtained the point estimators of stress-strength model when the stress and the strength both are independent Weibull random variables with common shape and different scale parameters based on RSS by using maximum likelihood (ML) and modified ML methodologies, Hassan et al. [17] used RSS for point and interval estimators of $P = P_r(Y < X)$ based on Gompertz distribution and MLES are compared by using MC simulation techniques. Hossein et al. [18] consider the RSS to estimate the parameters exponentiated pareto distribution and conclude that the estimator based on the ranked set sample have far better efficiency than the simple random sample at the same sample size. Akgul and Senoglu [19] constructed the asymptotic confidence interval for 'P' and obtained point and interval estimators for $P = P_r(Y < X)$ based on RSS, in addition the BCI for 'P' is constructed based on two distinct resampling methods.

In this paper we consider the point estimation of 'P' the stress-strength model, when the random stress and strength are two independent GIW random variables with different shape and scale parameters. A quick summary of the GIW distribution is given in section 2 and the point estimation of P using the maximum likelihood (ML) approach based on SRS is given in section 3. A brief explanation of RSS and its application in the point estimation is given in section 4. Monte Carlo simulation study is carried out in section 5 and a real life data study is performed for this model in section 6. Section 7 gives the concluding remarks for this study.

## 2. Preliminary

The GIW distribution is a continuous probability distribution which is proposed by de. Gosmao et al. [20]. It is an extended form of the Inverse Weibull distribution, introducing additional shape parameters to provide more flexibility in modelling. GIW has many applications in reliability, particularly in modelling the degradation of mechanical components such as pistons and crankshafts of diesel engines, as well as the breakdown of insulating fluid and in biological studies, where it is used to model a variety of failure characteristics such as infant mortality, useful life, and wear-out periods. Figure 1 and Figure 2 are showing the behaviour of probability density function and hazard rate function of GIW distribution respectively. The probability density function is positively skewed and the hazard rate function which is also known as failure rate function, during the initial phase, the hazard rate increases, indicating that the conditional probability of failure grows over time. This might represent a period where stress accumulation or wear-out effects dominate. However, after reaching a peak, the hazard rate begins to decrease, suggesting that units that have survived beyond a certain point have demonstrated their resilience and are less likely to fail immediately. This pattern can be observed in various real-world phenomena, such as certain mechanical systems or biological processes.

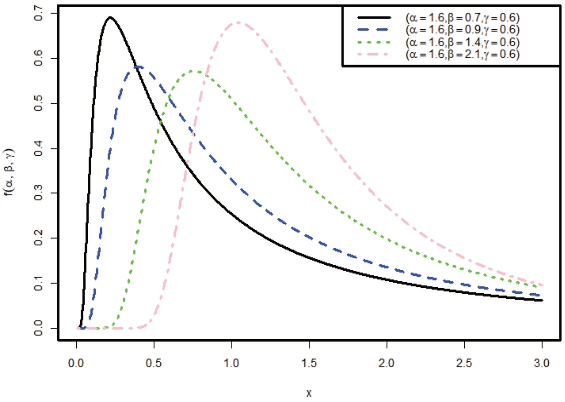
The probability density function (pdf) and cumulative distribution function (cdf) of GIW distribution are given respectively as

$$f(\,x, \alpha\, \beta, \gamma\,) = \gamma\beta\alpha^{\beta x^{-(\beta+1)}} \exp\left[-\gamma\left(\frac{\alpha}{x}\right)^{\beta}\right] \; ; \text{x, α, β, γ} > 0 \tag{2.1}$$

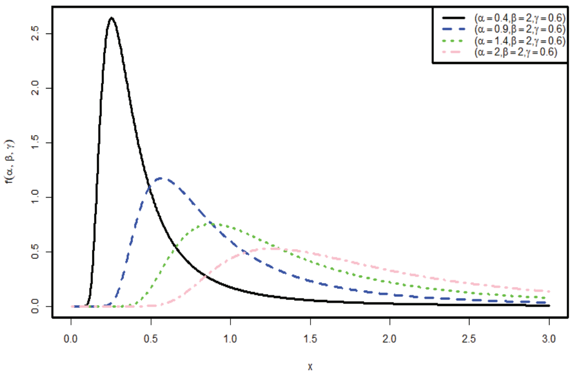
$$F(X) = \exp\left[-\gamma\left(\frac{\alpha}{x}\right)^{\beta}\right] \; ; \text{x, α, β, γ} > 0 \tag{2.2}$$

Hazard rate equation of GIW distribution given as follows-

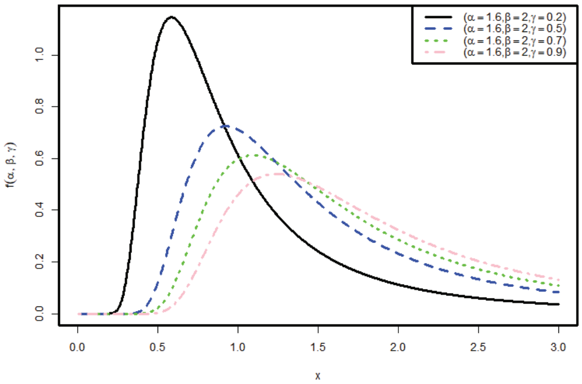
$$h(t) = \gamma\beta\alpha^{\beta}t^{-(\beta-1)}exp\left(-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right)\left[1 - exp\left(-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right)\right]^{-1} ; \text{t, α, β, γ} > 0 \tag{2.3}$$




*(a) For fixed α = 1.6 and γ = 0.6*

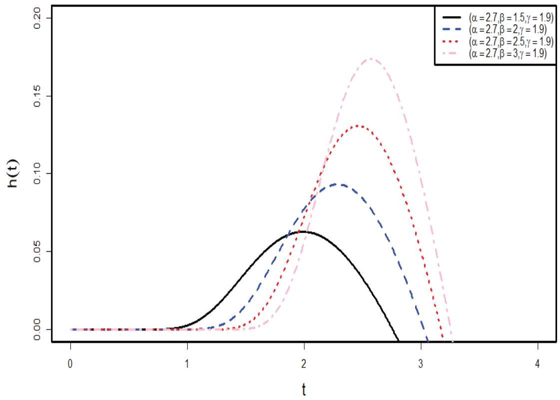



*(b) For fixed β = 2 and γ = 0.6*

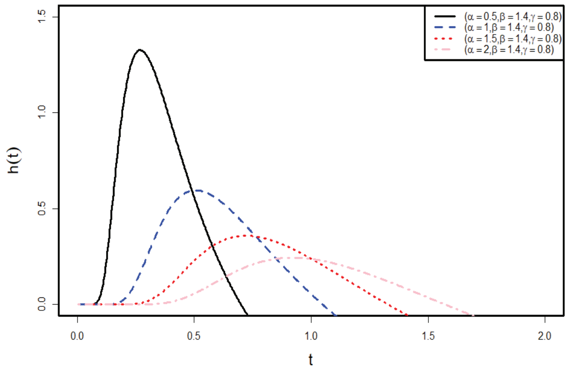



*(c) For fixed α = 1.6 and β = 2*

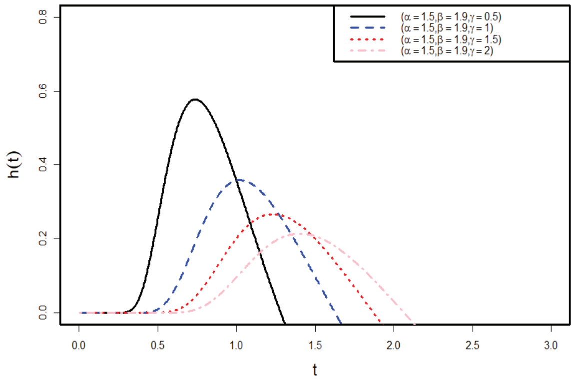
**Figure 1:** *Behaviour of pdf*




*(a) For fixed α = 2.7 and γ = 1.9*




*(b) For fixed β = 1.4 and γ = 0.8*




*(c) For fixed α = 1.5 and β = 1.9*

**Figure 2:** *Behaviour of Hazard rate*

## 3. Point Estimator for $P = P_r(Y < X)$ based on SRS

The pdf of GIW distribution is given by

$$f(x, \alpha\, \beta, \gamma) = \gamma\beta\alpha^{\beta x^{-(\beta+1)}} \exp\left[-\gamma\left(\frac{\alpha}{x}\right)^{\beta}\right] \; ; x, \alpha, \beta, \gamma > 0 \tag{3.1}$$

Let the rv's X and Y follow the GIW distribution given at (3.1) with the parameters $(\alpha, \beta, \gamma)$ and $(\theta, \mu, \chi)$.

**Theorem 3.1:** The MLE of $P = P_r(Y < X)$ is given by

$$P_{SRS}^{ML} = \frac{\bar{T}_Y}{(\bar{T}_X + \bar{T}_Y)}$$

$$where\ \bar{\Phi} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i{}^{-\beta} = \bar{T}_X\ (say), \qquad and\ \bar{\xi} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i{}^{-\mu} = \bar{T}_Y\ \text{(say)}$$

**Proof:** Let us consider the transformation $x^{-\beta} = \Phi$ in (3.1), we get

$$f(\Phi|\lambda) = \lambda exp[-\lambda\Phi] \ ; \Phi, \lambda > 0 \tag{3.2}$$

which is exponential distribution with parameter λ, where $\lambda = \gamma\alpha^{\beta}$

Let us considered Φ and ξ be two independent rv's which follows exponential distribution with parameters $\lambda_1$ and $\lambda_2$ respectively, where $x^{-\beta} = \Phi$ and $y^{-\mu} = \xi$

Thus for

$$P = P_r(\xi < \Phi)$$

$$P = \int_0^{\infty}\int_{\Phi=0}^{\xi} f(\Phi|\lambda_1)d\Phi f(\xi|\lambda_2)d\xi$$

$$P = \frac{\lambda_1}{(\lambda_1 + \lambda_2)} \tag{3.3}$$

If $\Phi_1, \Phi_2, \dots, \Phi_{n_1}$ *and* $\xi_1, \xi_2, \dots, \xi_{n_2}$ are two independent random samples of size $n_1$ *and* $n_2$ from the pdf's $f(\Phi|\lambda_1)$ and $f(\xi|\lambda_2)$ respectively then the joint pdf is given by

$$f(\Phi, \xi\,|\lambda_1, \lambda_2) = \lambda_1{}^{n_1}\,\lambda_2{}^{n_2} exp[-n_1\,\lambda_1\bar{\Phi} - n_2\,\lambda_2\bar{\xi}] \tag{3.4}$$

Taking likelihood function of (3.4) and derivatives w.r.to $\lambda_1$ *and* $\lambda_2$ and equating to zero, we get MLES of $\lambda_1$ *and* $\lambda_2$ respectively i.e.

$\lambda_1 = \frac{1}{\bar{\Phi}}$ and $\lambda_2 = \frac{1}{\bar{\xi}}$

The reliability function of P is

$$P_{SRS}^{ML} = \frac{\bar{\xi}}{(\bar{\Phi} + \bar{\xi})} \tag{3.5}$$

The equation (3.5) can be written as

$$P_{SRS}^{ML} = \frac{\bar{T}_Y}{(\bar{T}_X + \bar{T}_Y)}$$

## 4. Point Estimator for $\mathrm{P} = P_r(Y < X)$ based on RSS

In this section, we derive the ML estimator of P based on RSS. We first discuss about RSS, RSS is a specialized statistical sampling method designed to improve the efficiency and accuracy of estimating population parameters, particularly when dealing with populations that are highly heterogeneous or contain outliers. This sampling technique was introduced as an alternative to traditional sampling methods, such as SRS, in order to tackle the challenges posed by extreme values or skewed distributions in the population. A significant increase in precision can occasionally be obtained by using RSS as an alternative to SRS. In a work by G. A. McIntyre, it was first suggested in relation to evaluating herbage productivity. RSS procedures are given below:

I. Consider random sample $x_1, x_2, \dots, x_m$ by SRS each of size m.

II. To obtain k observations from a population
III. Then, rank order them according to a pre-defined attribute.
IV. The unit that is judged the smallest is included in your ranked set sample.
V. This first unit is called the first judgement order statistics and denoted by $X_{[1]}$.
VI. Then we repeat the same process k time, there for the sample size is obtained as $n = km$.
VII. For better understanding this entire process, see the following table:

| | | | | | |
|---|---|---|---|---|---|
| Cycle 1 | $X_{[1]1}$ | $X_{[2]1}$ | $X_{[3]1}$ | … | $X_{[k]1}$ |
| Cycle 2 | $X_{[1]2}$ | $X_{[2]2}$ | $X_{[3]2}$ | … | $X_{[k]2}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Cycle m | $X_{[1]m}$ | $X_{[2]m}$ | $X_{[3]m}$ | … | $X_{[k]m}$ |

## 4.1 The maximum likelihood estimator of $P = P_r(Y < X)$

Let $x_{ij}$; $i = 1,2,\dots,m_1$ and $j = 1,2,\dots,r_1$ denote the raked set sample of size $n_1 = r_1 m_1$ from GIW distribution with parameter $(\alpha,\beta,\gamma)$ where $m_1$ is the set size and $r_1$ is the number of cycles and $y_{kl}$; $k = 1, 2, \dots, m_2$ and $l = 1, 2, \dots, r_2$ denote the ranked set sample of size $n_2 = r_2 m_2$ from GIW distribution with parameter $(\theta,\mu,\chi)$ where $m_2$ is the set size and $r_2$ is the number of cycles. Then the pdf of $x_{ij}$ and $y_{kl}$

$$f_i(x_{ij}) = \frac{m_1\,!}{(i-1\,)\,!\,(m_1 - i\,)\,!}[F(x_{ij}\,)]^{i-1}[1-F(x_{ij})]^{m_1-i}\,f(x_{ij})$$

$$f_k(y_{kl}) = \frac{m_2\,!}{(k-1\,)\,!\,(m_2 - k\,)!}[F(\,y_{kl})]^{k-1}[1-F(y_{kl})]^{m_2-k}\,f(y_{kl})$$

Then the likelihood function based on RSS is given by

$$L = \prod_{i=1}^{r_1}\prod_{j=1}^{m_1} f_i\,(x_{ij}) \prod_{k=1}^{r_2}\prod_{l=1}^{m_2} g_k\,(y_{kl})$$

$$= \prod_{i=1}^{r_1}\left(\prod_{j=1}^{m_1}\frac{m_1!\,\left(\lambda_1\exp(-\lambda_1\Phi_{ij})\right)}{(i-1)!\,(m_1-i)!}\exp(-\lambda_1\Phi_{ij})\right)^{i-1}\left(-\exp(-\lambda_1\Phi_{ij})\right)^{m_1-i+1}$$

$$\prod_{k=1}^{r_2}\left(\prod_{l=1}^{m_2}\frac{m_2!\,\left(\lambda_2\exp(-\lambda_2\xi_{kl})\right)}{(k-1)!\,(m_2-k)!}\exp(-\lambda_2\xi_{kl})\right)^{k-1}[-\exp(-\lambda_2\xi_{kl})]^{m_2-k+1}$$

$$L = W\lambda^{n_1}\lambda^{n_2}\prod_{i=1}^{r_1}\left(\prod_{j=1}^{m_1}\exp\,(-\lambda_1\Phi_{ij})\right)^{i-1}\left[-\exp(-\lambda_1\Phi_{ij})\right]^{m_1-i+1}\exp(-\lambda_1\Phi_{ij})$$

$$\prod_{k=1}^{r_2}\left(\prod_{l=1}^{m_2}\exp\,(-\lambda_2\xi_{kl})\right)^{k-1}[1-\exp(-\lambda_2\xi_{kl})]^{m_2-k+1}\exp(-\lambda_2\xi_{kl}) \tag{4.1}$$

where

$$W = \prod_{i=1}^{r_1}\prod_{j=1}^{m_1}\frac{m_1}{(i-1)!\,(m_1-i)!}\prod_{k=1}^{r_2}\prod_{l=1}^{m_2}\frac{m_2}{(k-1)!\,(m_2-k)!}$$

Taking log

$$log\ L = log\ W + n_1 log\lambda_1 + n_2 log\lambda_2 + \sum_{i=1}^{r_1}\sum_{j=1}^{m_1}(i-1)\ log\ [exp(-\lambda_1\Phi_{ij}) - \lambda_1\sum_{i=1}^{r_1}\sum_{j=1}^{m_1}\Phi_{ij}$$
$$+ \lambda_1\sum_{i=1}^{r_1}\sum_{j=1}^{m_1}(m_1 - i + 1)\Phi_{ij} - \lambda_2\sum_{i=1}^{r_1}\sum_{j=1}^{m_1}\xi_{kl} + \sum_{k=1}^{r_2}\sum_{l=1}^{m_2}(k-1)\ log\ [exp(-\lambda_2\xi_{kl})$$
$$+ \lambda_2\sum_{k=1}^{r_2}\sum_{l=1}^{m_2}(m_2 - k + 1)\xi_{kl}$$

$$\frac{\partial logL}{\partial\lambda_1} = 0$$

Then,

$$\frac{n_1}{\lambda_1} - \sum_{i=1}^{r_1}\sum_{j=1}^{m_1}\frac{(i-1)\exp(-\lambda_1\Phi_{ij})\,\Phi_{ij}}{\exp(-\lambda_1\Phi_{ij})} - \sum_{i=1}^{r_1}\sum_{j=1}^{m_1}\Phi_{ij} + \sum_{i=1}^{r_1}\sum_{j=1}^{m_1}(m_1 - i + 1)\Phi_{ij} \tag{4.2}$$

$$\frac{\partial logL}{\partial\lambda_2} = 0$$

Then,

$$\frac{n_2}{\lambda_2} - \sum_{k=1}^{r_2}\sum_{l=1}^{m_2}\frac{(k-1)\exp(-\lambda_2\xi_{kl})\,\xi_{kl}}{\exp(-\lambda_2\xi_{kl})} - \sum_{k=1}^{r_2}\sum_{l=1}^{m_2}\Phi_{ij} + \sum_{k=1}^{r_2}\sum_{l=1}^{m_2}(m_1 - k + 1)\xi_{kl} \tag{4.3}$$

Using a numerical method, we ascertain the values of the ML estimators for $\lambda_1$ and $\lambda_2$ based on RSS shown by ${\lambda_1}_{RSS}^{ML}$ and ${\lambda_2}_{RSS}^{ML}$ and using the invariance property of the ML estimator, we get the maximum of reliability parameter P based on RSS as

$$P_{RSS}^{ML} = \frac{{\lambda_1}_{RSS}^{ML}}{{\lambda_1}_{RSS}^{ML} + {\lambda_2}_{RSS}^{ML}} \tag{4.4}$$

where, $\lambda_1 = \frac{1}{\Phi}$ and $\lambda_2 = \frac{1}{\xi}$

$$P_{RSS}^{ML} = \frac{\bar{\xi}_{RSS}^{ML}}{\bar{\Phi}_{RSS}^{ML} + \bar{\xi}_{RSS}^{ML}}$$

where,

$$\bar{\Phi}_{RSS}^{ML} = \frac{1}{n_1}\sum_{i=1}^{n_1}{x_i}^{-\beta} = \bar{T}_{x,RS\,S} \quad and \quad \bar{\xi}_{RSS}^{ML} = \frac{1}{n_2}\sum_{i=1}^{n_2}{y_i}^{-\mu} = \bar{T}_{y,RSS}$$

$$P_{SRS}^{ML} = \frac{\bar{T}_{y,RSS}}{(\bar{T}_{x,RSS} + \bar{T}_{y,RSS})}$$

## 5. Simulation Study

This section contains the simulation study that compares our suggested reliability estimator P based on RSS with the conventional reliability estimator of P based on SRS using the provided MSE and Bias values, $Bias(\hat{P}) = E(\hat{P} - P)$ and $MSE(P) = E(\hat{P} - P)^2$, respectively. The relative efficiency of the estimator of P is calculated as $= \frac{MSE(P_{MLE,SRS})}{MSE(P_{MLE,RSS})}$. If the value of relative efficiency is greater than one, it signifies that $\bar{P}_{SRS}$ is more efficient than the $\bar{P}_{RSS}$. Using the R programming language, all calculations were carried out. The following steps are used to explain the simulation study.

1. Generate 1000 simple random sample of $x_1, \dots, x_{n_1}$ and $y_1, \dots, y_{n_2}$ from Generalize Inverse Weibull distribution with the sample sizes $(n_1, n_2)$.
2. Generate 1000 random sample $x_{11}, \dots, x_{m_1 r_1}$ and $y_{11}, \dots, y_{m_2 r_2}$ from Generalize inverse Weibull distribution with set sizes $m_1 = m_2 = 3, 4, 5$ in case of number of cycles $r_1 = r_2 = 5$ and when $r_1 = r_2 = 10$ then set size $m_1 = m_2 = 2, 3, 4$.
3. Initially the parameter for X ~ GIW ($\alpha$, $\beta$, $\gamma$) distribution are taken as $\alpha = 2$, $\beta = 0.1$, $\gamma = 0.6$ and Y ~ GIW $(\theta, \mu, \chi)$, $\theta = 3$, $\mu = 0.2$, $\chi = 0.5$. After that we vary $\alpha = 2.5, 4$ and $\theta = 3.5, 5$ respectively and other parameters are fixed.
4. The MSEs relative efficiency and biased are calculated.

**Table1**: *Biases, MSES and RE of P under SRS and RSS when $\beta = 0.1$, $\gamma = 0.6$ and $\mu = 0.2$, $\chi = 0.5$ and $r_1 = r_2 = 5, 10$*

| | | | | SRS | | | RSS | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $r_1 = r_2 = 5$ | | | | |
| | $(n_1,n_2)$ | $(m_1,m_2)$ | $P_{True}$ | $\hat{P}_{SRS}$ | Bias | MSE | $\hat{P}_{RSS}$ | Bias | MSE | RE |
| $\alpha$=2, $\theta$=3 | (15,15) | (3,3) | 0.50797 | 0.51985 | 0.01187 | 0.000185 | 0.51981 | 0.01183 | 0.000165 | 1.12343 |
| | (15,20) | (3,4) | | 0.51953 | 0.01156 | 0.000172 | 0.51989 | 0.01191 | 0.000161 | 1.07190 |
| | (20,20) | (4,4) | | 0.51964 | 0.01167 | 0.000169 | 0.51966 | 0.01168 | 0.000152 | 1.11753 |
| | (20,25) | (4,5) | | 0.51965 | 0.01167 | 0.000165 | 0.51987 | 0.01189 | 0.000153 | 1.08270 |
| | (25,25) | (5,5) | | 0.51960 | 0.01163 | 0.000163 | 0.51996 | 0.01198 | 0.000154 | 1.06218 |
| $\alpha$=2.5, $\theta$=3.5 | (15,15) | (3,3) | 0.50584 | 0.51947 | 0.01362 | 0.000218 | 0.51995 | 0.01410 | 0.000215 | 1.01260 |
| | (15,20) | (3,4) | | 0.51946 | 0.01361 | 0.000212 | 0.51993 | 0.01408 | 0.000212 | 1.00436 |
| | (20,20) | (4,4) | | 0.51956 | 0.01371 | 0.000212 | 0.51965 | 0.01380 | 0.000201 | 1.05510 |
| | (20,25) | (4,5) | | 0.51956 | 0.01371 | 0.000209 | 0.51971 | 0.01386 | 0.000200 | 1.04142 |
| | (25,25) | (5,5) | | 0.51952 | 0.01367 | 0.000207 | 0.51982 | 0.01397 | 0.000202 | 1.02271 |
| $\alpha$=4, $\theta$=5 | (15,15) | (3,3) | 0.49976 | 0.51995 | 0.02018 | 0.000421 | 0.51995 | 0.02019 | 0.000416 | 1.01253 |
| | (15,20) | (3,4) | | 0.51994 | 0.02017 | 0.000418 | 0.51993 | 0.02017 | 0.000413 | 1.01343 |
| | (20,20) | (4,4) | | 0.51992 | 0.02015 | 0.000417 | 0.51998 | 0.02021 | 0.000414 | 1.00835 |
| | (20,25) | (4,5) | | 0.51992 | 0.02015 | 0.000416 | 0.51991 | 0.02014 | 0.000410 | 1.01551 |
| | (25,25) | (5,5) | | 0.51991 | 0.02014 | 0.000415 | 0.51998 | 0.02022 | 0.000412 | 1.00549 |
| | | | | | | $r_1 = r_2 = 10$ | | | | |
| | $(n_1,n_2)$ | $(m_1,m_2)$ | $P_{True}$ | $\hat{P}_{SRS}$ | Bias | MSE | $\hat{P}_{RSS}$ | Bias | MSE | RE |
| $\alpha$=2, $\theta$=3 | (20,20) | (2,2) | 0.50797 | 0.51986 | 0.01189 | 0.000175 | 0.51953 | 0.01155 | 0.000154 | 1.13637 |
| | (20,30) | (2,3) | | 0.51984 | 0.01186 | 0.000167 | 0.51984 | 0.01186 | 0.000156 | 1.07116 |
| | (30,30) | (3,3) | | 0.51984 | 0.01186 | 0.000163 | 0.51975 | 0.01177 | 0.000150 | 1.08497 |
| | (30,40) | (3,4) | | 0.51983 | 0.01186 | 0.000159 | 0.51980 | 0.01182 | 0.000149 | 1.06985 |
| | (40,40) | (4,4) | | 0.51970 | 0.01172 | 0.000153 | 0.51992 | 0.01194 | 0.000150 | 1.02107 |
| $\alpha$=2.5, $\theta$=3.5 | (20,20) | (2,2) | 0.50584 | 0.51974 | 0.01389 | 0.000217 | 0.51947 | 0.01362 | 0.00020 | 1.08420 |
| | (20,30) | (2,3) | | 0.51972 | 0.01387 | 0.000211 | 0.51972 | 0.01387 | 0.000203 | 1.03920 |
| | (30,30) | (3,3) | | 0.51972 | 0.01387 | 0.000208 | 0.51964 | 0.01379 | 0.000199 | 1.04873 |
| | (30,40) | (3,4) | | 0.51971 | 0.01386 | 0.000205 | 0.51969 | 0.01384 | 0.000198 | 1.03801 |
| | (40,40) | (4,4) | | 0.51960 | 0.01375 | 0.000201 | 0.51978 | 0.01394 | 0.000200 | 1.00415 |
| $\alpha$=4, $\theta$=5 | (20,20) | (2,2) | 0.49976 | 0.51992 | 0.02015 | 0.000417 | 0.51997 | 0.02020 | 0.000416 | 1.00415 |
| | (20,30) | (2,3) | | 0.51990 | 0.02013 | 0.000414 | 0.51989 | 0.02013 | 0.000410 | 1.00944 |
| | (30,30) | (3,3) | | 0.51990 | 0.02013 | 0.000413 | 0.51984 | 0.02007 | 0.000407 | 1.01444 |
| | (30,40) | (3,4) | | 0.51989 | 0.02013 | 0.000411 | 0.51987 | 0.02011 | 0.000407 | 1.00961 |
| | (40,40) | (4,4) | | 0.51982 | 0.02005 | 0.000408 | 0.51994 | 0.02018 | 0.000392 | 1.04033 |

The data presented in Table 1 consistently demonstrates that the relative efficiency exceeds unity, indicating the superior performance of ranked set sampling over simple random sampling in stress-strength reliability estimation.

## 6. Real data application

In order to confirm the results from earlier portions of the paper, we looked at two actual datasets in this section. We use two real-life data sets proposed by Efron B. [21]. The dataset includes patients from two groups who have head and neck cancer diseases. The survival times of 58 patients with radiotherapy-treated head and neck cancer are shown in the first dataset, whereas the survival times of 45 patients receiving chemotherapy plus radiation treatment are shown in the second dataset. In the context of stress–strength reliability, Yadav et al. [22] analysed these datasets and they showed that the Inverse Weibull distribution could be used to model these datasets. These datasets can also be useful in the case of generalized Weibull distribution. The first dataset of 58 patient is used for the strength variable X and the second dataset of 45 patients is used for stress variable Y in the stress-strength model $P = P_r(Y < X)$. The datasets are as follows:

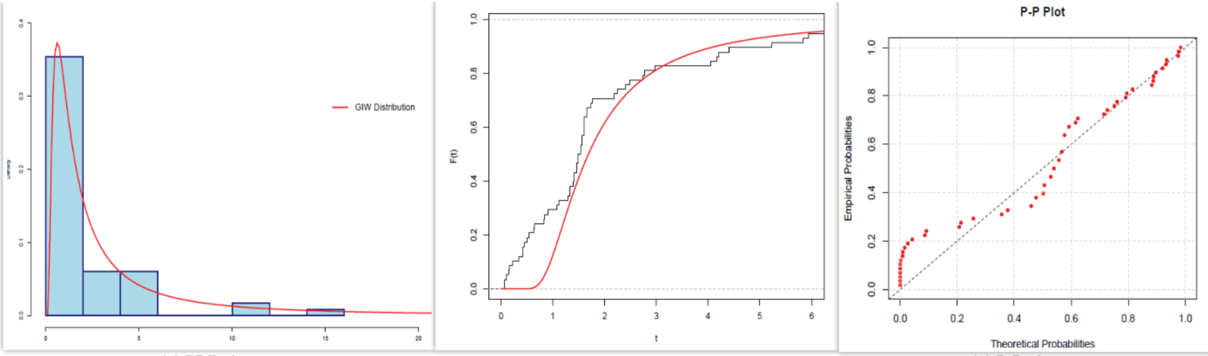
First data set of 58 patients

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 6.53 | 7 | 10.42 | 14.48 | 16.1 | 22.7 | 34 | 41.55 |
| 42 | 45.28 | 49.4 | 53.62 | 63 | 64 | 83 | 84 |
| 91 | 108 | 112 | 129 | 133 | 133 | 139 | 140 |
| 140 | 146 | 149 | 154 | 157 | 160 | 160 | 165 |
| 146 | 149 | 154 | 157 | 160 | 160 | 165 | 173 |
| 176 | 218 | 225 | 241 | 248 | 273 | 277 | 297 |
| 405 | 417 | 420 | 440 | 523 | 583 | 594 | 1101 |
| 1146 | 1417 | | | | | | |

Second data set of 45 patients

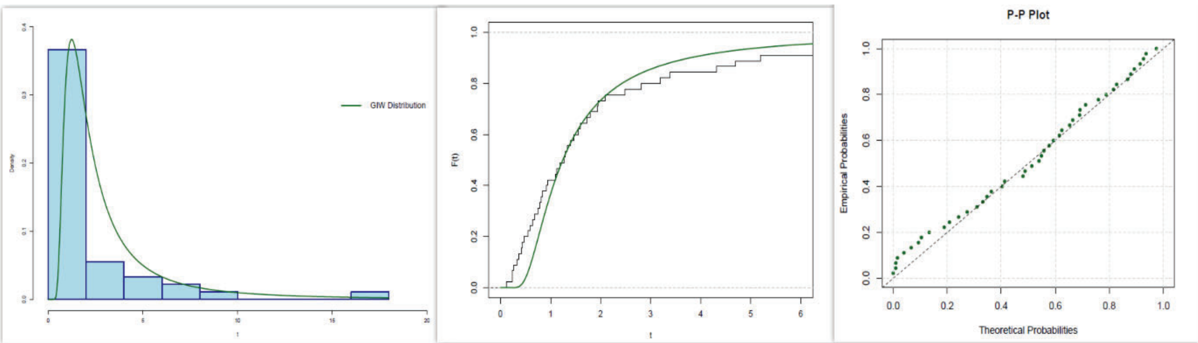
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 12.2 | 23.56 | 23.74 | 25.87 | 31.98 | 37 | 41.35 | 47.38 |
| 55.46 | 58.36 | 63.47 | 68.46 | 78.26 | 74.47 | 81 | 43 |
| 84 | 92 | 94 | 110 | 112 | 119 | 127 | 130 |
| 133 | 140 | 146 | 155 | 159 | 173 | 179 | 194 |
| 195 | 209 | 249 | 281 | 319 | 339 | 432 | 469 |
| 519 | 633 | 725 | 817 | 1776 | | | |

The first dataset of 58 patient is used for the strength variable X ~ GIW ($\alpha$, β, γ) and the second dataset of 45 patients is used for stress variable Y ~ GIW (θ, μ, χ) in the stress-strength model $P = P_r(Y < X)$. By using the iteration method in R software, the MLES of $\alpha$, β, γ and θ, μ, $\chi$ is comes out as $\hat{\alpha} = 2.9057$, $\hat{\beta} = 0.7859$, $\hat{\gamma} = 12.3257$ and $\hat{\theta} = 6.8548$, $\hat{\mu} = 1.0248$, $\hat{\chi} = 11.5366$. Now if we take these MLEs values of the parameters as the true value for these datasets then the stress-strength model $P = P_r(Y < X)$ from the Eq. (3.3) is $P = 0.25574$




(a) PDF plot (b) CDF plot (c) P-P plot

**Figure 3**: *The PDF, CDF and P-P Plots of the GIW distribution for First dataset*




(a) PDF plot (b) CDF plot (c) P-P plot

**Figure 4**: *The PDF, CDF and P-P Plots of the GIW distribution for First dataset*

Prior to delving into the core of our investigation, it is imperative to conduct a thorough examination of the salient characteristics of our dataset. To validate the robustness of our results, we employ a rigorous statistical methodology: the Kolmogorov-Smirnov (K-S) test, complemented by its associated P-value (P-V). This approach facilitates the quantification of the concordance between our empirical observations and theoretical expectations.

Our analysis yields promising results. For the initial dataset, we obtain a K-S distance of 0.31547 and a corresponding P-V of 0.42560. The secondary dataset exhibits comparable outcomes, with a K-S distance of 0.08889 and a P-V of 0.99520. These metrics provide substantial evidence supporting the close alignment of our model with the observed data.

To enhance comprehension and provide visual context, we have generated a series of graphical representations. These illustrations, presented in Figures 3 and 4, offer a comprehensive visualization of our statistical findings. They encompass probability-probability (PP) plots, as well as depictions of the estimated probability density function (PDF) and cumulative distribution function (CDF) for both datasets. These visual aids serve to corroborate and elucidate the numerical results of our analysis, thereby facilitating a more profound understanding of the data's underlying characteristics.

Now we draw 10 samples random sample of size 10 from each dataset and calculate the term $\bar{T}_x$ and $\bar{T}_y$ for each sample respectively. The simple random samples from each dataset are shown in Table 2 and Table 3, respectively.

**Table 2**: *Simple random samples from Data set 1*

| | | | | | | | | | | | $\bar{T}_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample 1 | 157 | 176 | 63 | 129 | 7 | 53.62 | 140 | 149 | 218 | 225 | 0.04257 |
| Sample 2 | 241 | 154 | 218 | 140 | 63 | 10.42 | 277 | 173 | 165 | 154 | 0.03312 |
| Sample 3 | 157 | 149 | 49.4 | 440 | 165 | 154 | 140 | 34 | 273 | 22.7 | 0.03118 |
| Sample 4 | 149 | 63 | 165 | 139 | 45.28 | 49.4 | 42 | 41.55 | 10.42 | 22.7 | 0.05443 |
| Sample 5 | 157 | 112 | 14.48 | 173 | 440 | 218 | 160 | 139 | 417 | 140 | 0.02745 |
| Sample 6 | 218 | 225 | 146 | 149 | 42 | 1417 | 176 | 7 | 297 | 53.62 | 0.04135 |
| Sample 7 | 165 | 157 | 133 | 63 | 173 | 420 | 45.28 | 112 | 176 | 583 | 0.02213 |
| Sample 8 | 225 | 140 | 165 | 277 | 149 | 14.48 | 91 | 129 | 108 | 157 | 0.03016 |
| Sample 9 | 157 | 64 | 157 | 63 | 154 | 146 | 165 | 241 | 16.1 | 133 | 0.03187 |
| Sample 10 | 140 | 22.7 | 417 | 176 | 139 | 149 | 146 | 41.55 | 583 | 241 | 0.02662 |

**Table 3:** *Simple random samples from Data set 2*

| | | | | | | | | | | | $\bar{T}_y$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample 1 | 1776 | 119 | 74.47 | 725 | 81 | 37 | 469 | 43 | 63.47 | 58.36 | 0.01097 |
| Sample 2 | 110 | 209 | 130 | 281 | 37 | 63.47 | 74.47 | 173 | 112 | 319 | 0.00889 |
| Sample 3 | 173 | 55.46 | 37 | 68.46 | 63.47 | 281 | 319 | 25.87 | 31.98 | 195 | 0.01481 |
| Sample 4 | 339 | 74.47 | 37 | 112 | 63.47 | 81 | 25.87 | 31.98 | 209 | 110 | 0.01492 |
| Sample 5 | 249 | 281 | 173 | 47.38 | 81 | 633 | 37 | 23.56 | 74.47 | 140 | 0.01256 |
| Sample 6 | 469 | 633 | 519 | 127 | 25.87 | 78.26 | 84 | 173 | 339 | 37 | 0.01019 |
| Sample 7 | 1776 | 94 | 817 | 23.56 | 469 | 130 | 43 | 146 | 173 | 68.46 | 0.01044 |
| Sample 8 | 127 | 110 | 78.26 | 249 | 209 | 41.35 | 94 | 43 | 432 | 155 | 0.00947 |
| Sample 9 | 110 | 159 | 23.74 | 194 | 31.98 | 633 | 1776 | 155 | 112 | 179 | 0.01061 |
| Sample 10 | 432 | 130 | 119 | 339 | 58.36 | 127 | 469 | 37 | 146 | 55.46 | 0.00902 |

In the next step, we draw 10 ranked set samples of size 10 from both the Data sets. To draw the ranked set sample of size n = 10, we take the set size m = 5 and run r = 2 cycles.

**Table 4** : *Ranked set samples from Data set 1*

| | | | | | | | | | | | $\bar{T}_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample1 | 7 | 112 | 157 | 165 | 405 | 10.42 | 108 | 139 | 160 | 225 | 0.05240 |
| Sample 2 | 139 | 225 | 297 | 165 | 583 | 10.42 | 64 | 417 | 112 | 440 | 0.03091 |
| Sample 3 | 7 | 165 | 165 | 160 | 173 | 112 | 63 | 49.4 | 146 | 160 | 0.04368 |
| Sample 4 | 42 | 63 | 176 | 160 | 149 | 6.53 | 133 | 154 | 133 | 277 | 0.04495 |
| Sample 5 | 22.7 | 7 | 165 | 405 | 1417 | 91 | 112 | 154 | 218 | 1417 | 0.04232 |
| Sample 6 | 91 | 64 | 146 | 160 | 1101 | 112 | 108 | 157 | 154 | 1101 | 0.02010 |
| Sample 7 | 84 | 139 | 405 | 149 | 1146 | 108 | 154 | 165 | 218 | 157 | 0.01795 |
| Sample 8 | 6.53 | 140 | 64 | 165 | 594 | 140 | 133 | 84 | 420 | 241 | 0.04069 |
| Sample 9 | 16.1 | 149 | 160 | 154 | 154 | 45.28 | 16.1 | 146 | 139 | 277 | 0.04040 |
| Sample 10 | 10.42 | 45.28 | 160 | 225 | 594 | 6.53 | 149 | 149 | 277 | 417 | 0.05364 |

**Table 5**: *Ranked set samples from Data set 2*

| | | | | | | | | | | | $\bar{T}_y$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample 1 | 43 | 55.46 | 195 | 319 | 249 | 23.56 | 41.35 | 155 | 130 | 173 | 0.01271 |
| Sample 2 | 25.87 | 94 | 58.36 | 159 | 469 | 63.47 | 68.46 | 146 | 194 | 817 | 0.01070 |
| Sample 3 | 12.2 | 63.47 | 173 | 146 | 469 | 41.35 | 155 | 112 | 155 | 633 | 0.01469 |
| Sample 4 | 63.47 | 112 | 119 | 339 | 249 | 25.87 | 84 | 81 | 81 | 725 | 0.01053 |
| Sample 5 | 47.38 | 55.46 | 195 | 249 | 817 | 41.35 | 92 | 155 | 281 | 1776 | 0.00855 |
| Sample 6 | 12.2 | 37 | 119 | 155 | 817 | 23.74 | 37 | 194 | 432 | 519 | 0.01877 |
| Sample 7 | 74.47 | 140 | 63.47 | 173 | 519 | 55.46 | 55.46 | 319 | 74.47 | 432 | 0.00887 |
| Sample 8 | 23.56 | 119 | 78.26 | 195 | 725 | 23.56 | 130 | 133 | 281 | 519 | 0.01213 |
| Sample 9 | 37 | 94 | 194 | 195 | 469 | 31.98 | 41.35 | 173 | 155 | 249 | 0.01100 |
| Sample 10 | 37 | 112 | 68.46 | 179 | 319 | 81 | 58.36 | 339 | 94 | 817 | 0.00930 |

To get 10 samples of size 10, we totally run 20 cycles. Each two subsequent cycles constitute one sample of size 10. The 20 cycles we performed to get the 10 ranked samples from population X and displayed in Table 4. Similarly, the 20 cycles from population Y are drawn and the 10 ranked set samples are shown in Table 5.

**Table 6**: *Bias, MSE and Relative efficiency of stress-strength model P in case of SRS and RSS*

| | SRS | | | | | RSS | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{T}_x$ | $\bar{T}_y$ | $\hat{P}_{SRS}$ | Bias | MSE | $\bar{T}_x$ | $\bar{T}_y$ | $\hat{P}_{RSS}$ | Bias | MSE | RE (%) |
| Sample 1 | 0.04257 | 0.01097 | 0.25278 | -0.00295 | 0.052 | 0.05240 | 0.01271 | 0.23154 | -0.0242 | 0.00872 | 5.92 |
| Sample 2 | 0.03312 | 0.00889 | | | | 0.03091 | 0.01070 | | | | |
| Sample 3 | 0.03118 | 0.01481 | | | | 0.04368 | 0.01469 | | | | |
| Sample 4 | 0.05443 | 0.01492 | | | | 0.04495 | 0.01053 | | | | |
| Sample 5 | 0.02745 | 0.01256 | | | | 0.04232 | 0.00855 | | | | |
| Sample 6 | 0.04135 | 0.01019 | | | | 0.02010 | 0.01877 | | | | |
| Sample 7 | 0.02213 | 0.01044 | | | | 0.01795 | 0.00887 | | | | |
| Sample 8 | 0.03016 | 0.00947 | | | | 0.04069 | 0.01213 | | | | |
| Sample 9 | 0.03187 | 0.01061 | | | | 0.04040 | 0.01100 | | | | |
| Sample10 | 0.02662 | 0.00902 | | | | 0.05364 | 0.00930 | | | | |

Based on Table 6, it can be inferred that the MSE of stress-strength model P under RSS conditions is lower than that of stress-strength model P under SRS circumstances. 5.92% of RSS is relative to SRS, or relative efficiency. Thus, in real-world scenarios, RSS techniques yield better outcomes than SRS strategies.

## 7. Conclusions

This work addresses the challenge of estimating the reliability function $P = P_r(Y < X)$, where X and Y are the independent random strength and stress variables from GIW distribution. The MLE of P is derived for SRS and RSS. Monte Carlo simulation study is performed to compare between point estimators of P in both cases SRS and RSS. The MLE of P based on RSS is more efficient results than the MLE of P based on the SRS. We further validate the advantages of RSS through an analysis of real-life data. Future research we shall explore the application of various type of ranked set sampling techniques in estimating the reliability models.

**Author Contributions:** All authors have equal contribution. All authors reviewed the results and approved the final version of manuscript.

**Funding:** This research received no Specific grant from any funding agency in the public, commercial, or not-for-profit-sectors.
**Conflict of Interest:** The authors declare no competing interests.

## References

[1] McIntyre, G.A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian journal of agricultural research*. ; 3(4), 385-390.

[2] Birnbaum, Z.W. (1956). On a use of the Mann-Whitney statistic. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, University of California Press. Jan; Vol. 3, pp. 13-18.

[3] Birnbaum, Z.W. and McCarty, R.C. (1958). A Distribution-Free Upper Confidence Bound for \Pr{Y<X\}, Based on Independent Samples of X and Y. *The Annals of Mathematical Statistics*; 29(2):558–62.

[4] Church J.D. and Harris, B. (1970). The estimation of reliability from stress-strength relationships. *Technometrics;* 12(1), 49-54.

[5] Chaturvedi, A. and Kumar, S. (1999). Further remarks on estimating the reliability function of exponential distribution under type I and type II censorings. *Brazilian Journal of Probability and Statistics*; 29-39.

[6] Kotz ,S., Lumelskii, Y., Pensky, M. (2003). The stress–strength model and its generalizations: theory and applications. World Scientific Publishing Co Pte Ltd.

[7] Kundu, D. and Gupta, R.D. (2005). Estimation of P [Y< X] for generalized exponential distribution. *Metrika*; 61, 291-308.

[8] Kundu, D. and Gupta, R.D. (2006). Estimation of P [Y< X] for Weibull distributions. *IEEE Trans. Reliab*; 55(2), 270-280.

[9] Raqab, M.Z. and Kundu, D. (2005). Coparison of different estimators of P [Y< X] for a scaled Burr type X distribution. *Commun. Statist. Simul. Computat; 34, 465-483.*

[10] Kundu, D. and Raqab, MZ. (2009). Estimation of R = P (Y < X) for three-parameter Weibull distribution. *Statistics & Probability Letters*; 79(17), 1839-1846.

[11] Krishnamoorthy, K., Mukherjee S., and Guo, H.(2007). Inference on reliability in two-parameter exponential stress-strength model. *Metrika; 65, 261-273.*

[12] Hassan, M.K.(2016). Estimation a Stress-Strength Model for P (Yr: n_1<Xk: n_2) Using the Lindley Distribution. *Revista Colombiana de estadistica*; 40(1), 105-121.

[13] Wang, B.X., Geng, Y. and Zhou, J.X.(2018). Inference for the generalized exponential stress-strength model. *Applied Mathematical Modelling*; 53, 267-275.

[14] Kayal, T., Tripathi, Y.M., Dey, S. and Wu, SJ.(2020). On estimating the reliability in a multicomponent stress-strength model based on Chen distribution. *Communications in Statistics-Theory and Methods*; 49(10), 2429-2447.

[15] Kumar, S., Chaturvedi, A.(2020). On a Generalization of the Positive Exponential Family of Distributions and the Estimation of Reliability Characteristics. *Statistica*; 80(1).

[16] Akgul, F.G., Acıtaş, Ş. and Şenoğlu, B.(2018). Inferences on stress–strength reliability based on ranked set sampling data in case of Lindley distribution. *Journal of Statistical Computation and Simulation*; 88(15), 3018-3032.

[17] Hassan, M.K., Alohali, M.I. and Alojail, F.A.(2021). Point and Interval Estimators of R= P [Y< X] Based on Gompertz Distribution and Ranked Set Sampling Data. *Thailand Statistician*; 19(4), 784-796.

[18] Khamnei, H.J., Meidute-Kavaliauskiene, I., Fathi, M., Valackienė, A. and Ghorbani, S.(2022). Parameter Estimation of the Exponentiated Pareto Distribution Using Ranked Set Sampling and Simple Random Sampling. *Axioms*; 11(6), 293.

[19] Akgul, F.G., Şenoğlu, B.(2022). Inferences for stress–strength reliability of Burr Type X distributions based on ranked set sampling. *Communications in Statistics-Simulation and Computation*; 51(6), 3324-3340.

[20] De Gusmao, F.R., Ortega, E.M. and Cordeiro, G.M.(2011) The generalized inverse Weibull distribution. *Statistical Papers*; 52, 591-619.

[21] Efron, B.(1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association*; 83(402), 414-425.

[22] Singh, S.K., Singh, U. and Yadav, AS.(2015). Reliability estimation and prediction for extension of exponential distribution using informative and non-informative priors. *International Journal of System Assurance Engineering and Management*; *6, 466-478.*